SI

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/007,886	11/05/2001	Changxue Ma	CML00075H	3158

22917 7590 09/22/2004

MOTOROLA, INC.
1303 EAST ALGONQUIN ROAD
IL01/3RD
SCHAUMBURG, IL 60196

EXAMINER	
ALBERTALLI, BRIAN LOUIS	
ART UNIT	PAPER NUMBER
2655	

DATE MAILED: 09/22/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)
	10/007,886	MA ET AL.
	Examiner	Art Unit
	Brian L Albertalli	2655

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on ____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-21* is/are pending in the application.

4a) Of the above claim(s) ____ is/are withdrawn from consideration.

5)☐ Claim(s) ____ is/are allowed.

6)☒ Claim(s) *1-21* is/are rejected.

7)☒ Claim(s) *12 and 13* is/are objected to.

8)☐ Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *05 November 2001* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☒ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. ____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *11/5/01, 6/2/04*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. ____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: ____.

DETAILED ACTION

Oath/Declaration

1. The oath or declaration is defective. A new oath or declaration in compliance with 37 CFR 1.67(a) identifying this application by application number and filing date is required. See MPEP §§ 602.01 and 602.02.

The oath or declaration is defective because: Changxue Ma's signature is not dated.

Specification

2. The disclosure is objected to because of the following informalities: on page 18, line 16, "202" should be –204--.

Appropriate correction is required.

Claim Objections

3. Claim 12 is objected to because of the following informalities: in the last line (line 18) the semicolon ";" should be a period --.--.

Claim 13 is objected to because of the following informalities: in line 14 of the claim, "a" should be deleted.

Appropriate correction is required.

Claim Rejections - 35 USC § 102

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

5. Claims 1 and 10-13 are rejected under 35 U.S.C. 102(b) as being anticipated by Narayanan et al. (U.S. Patent 6,076,057).

In regard to claims 1 and 13, Narayanan et al. discloses a method and a computer readable medium storing program instructions (software) for performing automatic speech recognition in a variable background noise environment, the method comprising the steps of:

Processing a first portion of an audio signal to obtain a first characterization of the first portion of the audio signal (Fig. 5, step 510, input utterance, column 5, lines 34-37; utterance is input by A/D converter, Fig. 2, 210 and characterized by the feature extraction unit 220, column 3, lines 7-35);

Comparing the first characterization to a set of reference characterizations to determine a particular reference characterization among the set of reference characterizations that most closely matches the first characterization (step 550, competing strings are optimally decoded to find optimum segmentation, column 5, lines 42-44; comparing is performed by pattern matching processor 310, column 4, line 46 through column 5, line 20);

Updating the particular reference characterization so that the particular reference characterization more closely resembles the first characterization (step 560 Hidden Markov Models, HMM's, are adapted, column 5, lines 44-46).

In regard to claim 13, Narayanan et al. additionally discloses processing one or more additional portions of the audio signal (the entire utterance) to obtain one or more additional characterizations that characterize the one or more additional portions of the audio signal;

comparing the one or more additional characterizations to the set of reference characterization to find reference characterizations among the set of reference characterizations that most closely matches the one or more additional characterizations (recognition is performed on the entire utterance using the newly adapted HMM's, column 5, lines 47-51).

In regard to claim 10, Narayanan et al. discloses an automated speech recognition system comprising:

an audio signal input for inputting an audio signal that includes speech and background sounds (Fig. 1, transducer 105, column 2, lines 57-58);

a feature extractor coupled to the audio signal input for receiving the audio signal and outputting characterizations of a sequence of segments of the audio signal (feature extraction unit 220, column 3, lines 22-35);

a model coupled to the feature extractor, wherein the model includes a plurality of states to which characterization of the sequence of segments are applied for evaluating a posteriori probabilities that one or more of the plurality of states occurred (acoustic model unit 320 stores HMM model, column 3, line 57 through column 4, line 34);

a search engine coupled to model for finding one or more high probability sequences of the plurality of states of the model (pattern matching processor 310, column 4, lines 46-65);

a detector for detecting a specific state of the audio signal and outputting a predetermined signal when the specific state is detected (recognizer performs speech silence segmentation, column 5, lines 52-58);

and a comparer and updater coupled to the detector for receiving the predetermined signal and in response thereto updating the model so that it more closely models one or more characterizations output by the feature extractor that correspond to the specific state (pattern matching processor 310 performs the steps of updating the Hidden Markov Models, column 5, lines 44-46).

In regard to claim 11, Narayanan et al. discloses the feature extractor outputs characterizations for each of a succession of frames that include feature vectors that include cepstral coefficients (Cepstral analysis is employed to obtain the features of the speech signal, column 3, lines 22-26);

the model comprises a hidden markov model that includes a plurality of emitting states and multi component Gaussian mixtures that give the a posteriori probability that a given feature vector is attributable to a given emitting state (column 4, lines 19-34, and Fig. 4);

the detector detects an absence of speech sounds by comparing a function of one more cepstral coefficients to a threshold (silence regions are identified based on

signal power, column 5, lines 26-29. The signal power inherently must be compared to some threshold to make a speech/silence decision); and

the comparer and updater determines a mean of a multi component Gaussian mixture associated with background sounds that is closest to a feature vector that characterizes the audio signal during the absence of speech sounds, and updates the mean so that it is closer to the feature vector that characterizes the audio signal during the absence of speech sounds (the optimal decoding of the silence regions using background HMM's is performed, the HMM's of the silence regions are then adapted, column 3, lines 37-39 and lines 44-46, column 6, lines 14-18).

In regard to claim 12, Narayanan et al. discloses an automated speech recognition system comprising:

an audio input for inputting an audio signal (transducer 105, column 2, lines 57-58);

an analog to digital converter coupled to the audio input for sampling the audio signal and outputting a discretized audio signal (A/D converter 210, column 3, lines 7-9); and

a microprocessor coupled to the analog to digital converter for receiving the discretized audio signal and executing a program for performing automated speech recognition (digital signal processor, column 2, lines 49-52), the program comprising programming instructions for:

processing a first portion of an audio signal to obtain a first characterization of the first portion of the audio signal (Fig. 5, step 510, input utterance, column 5, lines 34-37; utterance is input by A/D converter, Fig. 2, 210 and characterized by the feature extraction unit 220, column 3, lines 7-35);

comparing the first characterization to a set of reference characterizations to determine a particular reference characterization among the set of reference characterizations that most closely matches the first characterization (step 550, competing strings are optimally decoded to find optimum segmentation, column 5, lines 42-44; comparing is performed by pattern matching processor 310, column 4, line 46 through column 5, line 20); and

updating the particular reference characterization so that the particular reference characterization more closely resembles the first characterization (step 560 Hidden Markov Models, HMM's, are adapted, column 5, lines 44-46).

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

7. Claims 2-7 and 14-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Narayanan et al., in view of Wang (U.S. Patent 5,594,834).

In regard to claims 2 and 14, Narayanan et al. discloses detecting a pause in the speech (step 520, input utterance 510 is split into speech and silence regions, column 5, lines 36-37); and

In response to the step of detecting, performing the step of processing the first portion of the audio signal wherein the first portion of the audio signal is included in the pause (silence region is detected and used to adapt the silence models, column 5, lines 30-33).

Narayanan et al. does not disclose that the pause is an inter sentence pause.

Wang discloses that continuously spoken speech only contains pauses at "natural" points, such as the end of a sentence (column 4, lines 3-17).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Narayanan et al. to detect inter sentence pauses, so the user would be able to speak continuously to the recognizer and would not have to un-naturally pause between each sentence.

In regard to claims 3 and 15, Narayanan et al. discloses the step of processing the first portion of the audio signal to obtain a first characterization includes a sub-step of:

Processing the first portion of the audio signal to obtain a first set of numbers that characterize the first portion of the audio signal (audio signal is represented by a set of vectors defining the parameters of the HMM, column 3, line 57 through column 4, line 34); and

The step of comparing the first characterization to a set of reference characterizations comprises the sub-steps of:

Comparing the first set numbers to a plurality of reference sets of numbers to determining a particular set of reference numbers that most closely matches the first set of numbers (pattern matching processor 310 searches the network of HMM's stored in acoustic unit 320 to find the most-likely match, column 4, lines 46-65).

In regard to claims 4 and 16, Narayanan et al. discloses the step of updating the reference characterization comprises the sub-steps of:

Replacing each number in the particular set of numbers with a weighted average of the number and a corresponding number in the first set of numbers. Narayanan et al. discloses the HMM's are adapted using a discriminative training algorithm (column 4, line 66 through column 5, line 20 and column 6, lines 22-25). Unsupervised adaptation using a gradient descent algorithm updates the parameters of the HMM model using a model dependent weighting term.

In regard to claim 5 and 17, Narayanan et al. does not disclose comparing the first characterization to a set of reference characterizations comprises the sub-steps of:

taking a dot product between the first set of numbers and each of the plurality of reference sets of numbers.

Official notice is taken that it is notoriously well known and recognized in the art to calculate the dot product between two vectors to determine the similarity between two vectors.

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Narayanan et al. and Wang to take the dot product between the first set of numbers and each of the plurality of reference sets of numbers since the dot product is a simple calculation that can be calculated quickly, thereby decreasing the amount of time needed to update a set of reference numbers.

In regard to claims 6 and 18, Narayanan et al. discloses the plurality of reference sets of numbers characterize a plurality of types of non speech audio (more than one background HMM is used, column 5, lines 37-39).

In regard to claims 7 and 19, Narayanan et al. discloses the plurality of reference sets of numbers are means of components of Gaussian mixtures that characterize the probability of an underlying state of a hidden markov model of the audio signal, given the first set of numbers (column 3, line 57 through column 4, line 34, particularly column 4, line 30).

8. Claims 8-9 and 20-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Narayanan et al., in view of Wang, and in further view of Laurila et al. (U.S. Patent 6,772,117).

In regard to claims 8 and 20, Narayanan et al. discloses the step of processing the first portion of the audio signal to obtain the first characterization of the first portion of the audio signal comprises the sub-steps of:

a) time domain sampling the audio signal to obtain a discretized representation of the audio signal that includes a sequence of samples (A/D converter 210 transforms analog waveform signals into digital signals, column 3, lines 7-9);

b) time domain filtering the sequence of samples to obtain a filtered sequence of samples (anti-aliasing filter, column 3, lines 9-10);

c) applying a window function to successive subsets of the filtered sequence of samples to obtain a sequence of frames of windowed filtered samples (column 3, lines 30-33);

Narayanan et al. further suggests several different techniques to extract features from the windows of the speech signal. One of the techniques is Cepstral analysis.

Narayanan et al. is silent as to the steps taken to perform Cepstral analysis to extract features from the windows of the input signal.

Laurila et al. discloses a method of extracting Cepstral features from an input signal. The method includes the steps described above, and further includes:

d) transforming each of the frames of windowed filtered samples to a frequency domain to obtain a plurality of frequency components (Fig. 2, FFT 23, column 3, lines 41-46);

e) taking a plurality of weighted sums of the plurality of frequency components to obtain a plurality of bandpass filtered outputs (Mel windowing block 24, column 3, lines 46-48);

f) taking the log of the magnitude of each of the bandpass filtered outputs to obtain a plurality of log magnitude bandpass filtered outputs (25, column 3, lines 59-60); and

g) transforming the plurality of log magnitude bandpass filtered outputs to a time domain to obtain at least a subset of the first set of numbers (DCT 26, column 3, lines 60-63).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Naryanan et al. to extract the features of the windowed input signal in the manner described by Laurila et al., since, as is well known in the art, Mel-Frequency Cepstral Coefficients provide a compact means to represent speech.

In regard to claims 9 and 21, the combination of Naryanan et al., Wang, and Laurila et al., discloses in Laurila et al. repeating sub-steps (a) through (g) for two portions of the audio signal to obtain two sets of numbers (successive output vectors of discrete cosine transformation block 26); and

taking the difference between corresponding numbers in the two sets of numbers to obtain at least a subset of the first set of numbers (column 3, line 63 through column 4, line 5).

Conclusion

9. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Reichl et al. (*Discriminative Training for Continuous Speech Recognition*) discloses unsupervised adaptation using a gradient descent algorithm updates the parameters of the HMM model using a model dependent weighting term. Campbell et al. (U.S. Patent 6,131,089) discloses a speech recognition system that calculates a dot product between feature vectors. Winn (U.S. Patent 6,108,610) discloses a method of adapting noise estimates during pauses in speech. Downey (U.S. Patent 6,078,884) discloses a system that generates a noise model for each noise portion of the input signal. Tzirkel-Hanckock (U.S. Patent 5,960,395) discloses an additional feature extraction preprocessor that calculates Cepstral coefficients. Wu et al. (U.S. Patent 6,778,959) discloses a system that updates a plurality of noise models.

10. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brian L Albertalli whose telephone number is (703) 305-1817. The examiner can normally be reached on Monday - Friday, 8:30 AM - 5:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (703) 305-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

BLA 9/16/04

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER